

CLAIMS:

What is claimed is:

- 5 1. A method comprising:
- filtering a frame of a first speech signal to obtain a residual signal frame and a
 set of vocal tract model parameters, wherein the frame of the first speech signal and
 the residual signal frame contain a same fixed number of samples;
- determining from the residual signal frame at least one pitch cycle within the
10 residual frame;
- applying a transformation function to the residual frame to obtain a modified
 residual frame, wherein the modified residual frame contains an integer number of
 pitch cycles; and
- synthesizing a second speech signal from the modified residual frame and the
15 set of vocal tract model parameters, whereby the second speech signal is a
 pitch-compensated speech signal.
2. The method of claim 1, wherein the integer number of pitch cycles is a
 predetermined integer number of pitch cycles.
- 20 3. The method of claim 2, wherein the integer number of pitch cycles is
 predetermined to be one.
4. The method of claim 1, wherein the transformation function changes a time
25 scale of a residual signal represented by the residual signal frame.

5. The method of claim 4, wherein the transformation function changes the time scale of the residual signal by performing operations that include:

selecting a set of samples from the residual signal, wherein the set of samples is a consecutive sequence of samples taken from the residual signal, such that the set
5 of samples corresponds to a contiguous interval of time in the residual signal;

performing linear interpolation between samples in the first set of samples so as to model the residual signal over said contiguous interval of time as a piecewise linear function; and

generating the modified residual signal by generating a new sequence of
10 samples from the piecewise linear function such that the cardinality of the new sequence of samples is equal to the same fixed number of samples as contained in the residual signal frame.

6. The method of claim 4, wherein the transformation function changes the time
15 scale of the residual signal by performing a non-linear time warping operation on an interval of the residual signal so as to find a correspondence between samples from the interval of the residual signal and samples in a reference signal.

7. The method of claim 6, wherein the non-linear time warping operation is
20 performed according to a dynamic time warping algorithm.

8. The method of claim 1, wherein the transformation function generates a modified residual signal frame from the residual signal frame by performing operations that include:

25 mapping to zero a first subset of samples from a residual signal represented by the residual signal frame; and

mapping a second subset of samples from the residual signal to their identical values.

9. The method of claim 1, further comprising:
cyclically shifting samples in the modified residual signal frame so as to
normalize a phase of the modified residual signal frame.
- 5
10. The method of claim 1, further comprising:
feeding the modified residual signal frame to at least one of speech
recognition software and speaker recognition software.
- 10 11. A computer program product in a computer-readable medium comprising
functional descriptive material that, when executed by a computer, causes the
computer to perform acts that include:
filtering a frame of a first speech signal to obtain a residual signal frame and a
set of vocal tract model parameters, wherein the frame of the first speech signal and
15 the residual signal frame contain a same fixed number of samples;
determining from the residual signal frame at least one pitch cycle within the
residual frame;
applying a transformation function to the residual frame to obtain a modified
residual frame, wherein the modified residual frame contains an integer number of
20 pitch cycles; and
synthesizing a second speech signal from the modified residual frame and the
set of vocal tract model parameters, whereby the second speech signal is a
pitch-compensated speech signal.
- 25 12. The computer program product of claim 11, wherein the integer number of
pitch cycles is a predetermined integer number of pitch cycles.

13. The computer program product of claim 12, wherein the integer number of pitch cycles is predetermined to be one.
- 5 14. The computer program product of claim 11, wherein the transformation function changes a time scale of a residual signal represented by the residual signal frame.
15. The computer program product of claim 14, wherein the transformation
10 function changes the time scale of the residual signal by performing operations that include:
- selecting a set of samples from the residual signal, wherein the set of samples is a consecutive sequence of samples taken from the residual signal, such that the set of samples corresponds to a contiguous interval of time in the residual signal;
 - 15 performing linear interpolation between samples in the first set of samples so as to model the residual signal over said contiguous interval of time as a piecewise linear function; and
 - generating the modified residual signal by generating a new sequence of samples from the piecewise linear function such that the cardinality of the new
20 sequence of samples is equal to the same fixed number of samples as contained in the residual signal frame.
16. The computer program product of claim 14, wherein the transformation
25 function changes the time scale of the residual signal by performing a non-linear time warping operation on an interval of the residual signal so as to find a correspondence between samples from the interval of the residual signal and samples in a reference signal.

17. The computer program product of claim 16, wherein the non-linear time warping operation is performed according to a dynamic time warping algorithm.

18. The computer program product of claim 11, wherein the transformation
5 function generates a modified residual signal frame from the residual signal frame by performing operations that include:

mapping to zero a first subset of samples from a residual signal represented by the residual signal frame; and

mapping a second subset of samples from the residual signal to their identical
10 values.

19. The computer program product of claim 11, comprising additional functional descriptive material that, when executed by the computer, causes the computer to perform additional acts that include:

15 cyclically shifting samples in the modified residual signal frame so as to normalize a phase of the modified residual signal frame.

20. The computer program product of claim 11, comprising additional functional descriptive material that, when executed by the computer, causes the computer to
20 perform additional acts that include:

feeding the modified residual signal frame to at least one of speech recognition software and speaker recognition software.

21. A data processing system comprising:
25 means for filtering a frame of a first speech signal to obtain a residual signal frame and a set of vocal tract model parameters, wherein the frame of the first speech signal and the residual signal frame contain a same fixed number of samples;
means for determining from the residual signal frame at least one pitch cycle

within the residual frame;

means for applying a transformation function to the residual frame to obtain a modified residual frame, wherein the modified residual frame contains an integer number of pitch cycles; and

- 5 means for synthesizing a second speech signal from the modified residual frame and the set of vocal tract model parameters,
whereby the second speech signal is a pitch-compensated speech signal.

22. The data processing system of claim 21, wherein the integer number of pitch
10 cycles is a predetermined integer number of pitch cycles.

23. The data processing system of claim 22, wherein the integer number of pitch cycles is predetermined to be one.

15 24. The data processing system of claim 21, wherein the transformation function changes a time scale of a residual signal represented by the residual signal frame.

25. The data processing system of claim 24, wherein the transformation function changes the time scale of the residual signal by performing operations that include:

20 selecting a set of samples from the residual signal, wherein the set of samples is a consecutive sequence of samples taken from the residual signal, such that the set of samples corresponds to a contiguous interval of time in the residual signal;

performing linear interpolation between samples in the first set of samples so as to model the residual signal over said contiguous interval of time as a piecewise

25 linear function; and

generating the modified residual signal by generating a new sequence of samples from the piecewise linear function such that the cardinality of the new sequence of samples is equal to the same fixed number of samples as contained in the residual signal frame.

5

26. The data processing system of claim 24, wherein the transformation function changes the time scale of the residual signal by performing a non-linear time warping operation on an interval of the residual signal so as to find a correspondence between samples from the interval of the residual signal and samples in a reference signal.

10

27. The data processing system of claim 26, wherein the non-linear time warping operation is performed according to a dynamic time warping algorithm.

28. The data processing system of claim 21, wherein the transformation function generates a modified residual signal frame from the residual signal frame by performing operations that include:

15

mapping to zero a first subset of samples from a residual signal represented by the residual signal frame; and

20

mapping a second subset of samples from the residual signal to their identical values.

29. The data processing system of claim 21, further comprising:
means for cyclically shifting samples in the modified residual signal frame so as to normalize a phase of the modified residual signal frame.

25

30. The data processing system of claim 21, further comprising:
means for feeding the modified residual signal frame to at least one of speech recognition software and speaker recognition software.